A parsimonious method for generating syntactic structure

D.P. Medeiros, University of Arizona

November 2020

Abstract

This paper proposes a reformulation of (External) Merge, replacing the now-standard set-based version of the structure-building operation. In the present account, *n*-ary branching ordered trees are freely generated, followed by hierarchization and then linearization, which proceed by post-order traversal and pre-order traversal, respectively. This generative system accounts for constraints on word order in several domains: all and only cross-linguistically attested neutral orders in these domains are base-generated, with a systematic gap in generative capacity corresponding to the typological gap of impossible neutral orders. The framework unifies Universal 20 (Greenberg 1963, Cinque 2005) effects and the Final Over Final Condition (Holmberg 2000, Sheehan *et al* 2017) as consequences of this generative gap. As a consequence of base-generating neutral displacement, the cut for the duality of semantics is repositioned, properly discriminating the full range of neutral orders from genuine discourse/scope-related displacement. Beyond delimiting possible orders, the account assigns them appropriate bracketed structures. Finally, the system provides a straightforward account of attested structures containing cross-serial dependencies, including bounded relations as in English Affix-Hopping (Chomsky 1957), and unbounded relations as in Dutch (Bresnan *et al* 1982) and Swiss German (Shieber 1985) cross-serial subject-verb dependencies.

Keywords: Merge, Final-Over-Final Condition, Universal 20, Cross-serial dependencies

1. Introduction

Chomsky describes the discrete infinite character of human syntax in terms of an abstract operation Merge, a minimal mechanism to recursively generate complex expressions. Merge takes as input lexical elements or already-built syntactic objects (built by other Merge operations), and produces as output a structured expression containing its inputs. There are various ways of working out the details, but in broad strokes something like Merge seems indispensable.

Attention has largely focused on the implementation of Merge as bare set formation, which readily provides for a rich theory of syntactic structure. That implementation, whatever its successes and *a priori* appeal[1], is certainly not the only conceivable possibility; some caution is warranted in ascertaining the computational description of this novel biological capacity. If another reasonable implementation of Merge provides a different characterization of syntactic phenomenology, the alternatives should be evaluated by their empirical successes in addition to their conceptual properties.[2]

In recent years, Chomsky has highlighted the need for syntactic theories to provide a basis for the duality of semantics: the existence, in natural language expressions, of two layers of meaning. One layer of meaning is the information-neutral thematic structure, encompassing predicate-argument structure and selectional relations. Another layer of meaning concerns operator-variable structure, topic and focus, and the like. This empirical cut is, ideally, to be tied to some syntactically defined distinction; one suggestion is that it reflects a distinction in how Merge applies. If Merge joins two previously unconnected syntactic

[1] There are several reasons for preferring a set-based implementation for Merge. One reason is the same logic that drove mathematicians to look to set theory as an axiomatic basis for the rest of mathematics: it is maximally conceptually sparse. Another motivation for supposing that syntactic representations are unordered is that semantic composition seems to be describable in terms that eschew linear ordering, which looks instead like a property that is relevant only within the sensorimotor interface. But little enough is known that we should be careful here; see fn. 7.

[2] Compare, for example, the choice among the set of real numbers and the set of complex numbers for modeling physical phenomena. The reals seem conceptually inevitable and simpler than the complex numbers, of which they form a strict subset. However, the complex numbers provide a better basis for understanding physical phenomena like electromagnetism, and with their greater complexity comes greater mathematical beauty (for example, in the context of the Fundamental Theroem of Algebra).

objects, it is External Merge (EM). Where Merge applies to an object and one of its subparts, we have Internal Merge (IM).

> "The two types of Merge correlate well with the duality of semantics that has been studied from various points of view over the years. EM yields generalized argument structure, and IM all other semantic properties: discourse-related and scopal properties. The correlation is close, and might turn out to be perfect if enough were understood." (Chomsky 2007: 10)

It has been proposed that External Merge proceeds according to a common order of operations across languages, at least with respect to functional structure. The assumption of a universal ordering of External Merge is an essential component of the so-called cartographic program (Rizzi 1997, Cinque 1999), inserting elements from universal hierarchies into a bottom-up derivation. Internal Merge operations interleave with External Merge, (ultimately) yielding displacement. If External Merge applies in a common order, and syntactic structures are linearized in the same way across languages (Kayne 1994), it follows at once that Internal Merge must be involved in the derivation of word order variation.

But languages plainly vary in word order, even in information-neutral contexts. Information-neutral contexts, by definition, do not involve discourse or scopal properties. So what drives Internal Merge in the derivation of neutral orders?[3] Moreover, how can we explain the apparently complex constraints on possible and impossible neutral word orders? The problem is particularly acute for orders that contain crossing dependencies, which clearly exceed the context-free generative power of set-based External Merge.

2. Generating Universal 20

As a concrete example, consider possible and impossible neutral orders in the noun phrase, as first described in Greenberg's Universal 20.

> "When any or all of the items (demonstrative, numeral, and descriptive adjective) precede the noun, they are always found in that order. If they follow, the order is either the same or its exact opposite." (Greenberg 1963: 87)

According to Cinque's (2005) analysis, 14 of the 24 logically possible orders of these four elements are attested. Cinque shows that this pattern can be succinctly described within the Internal-and-External-Merge framework by assuming a universal, cartographically structured underlying base, built by a uniform sequence of External Merge operations, together with phrasal movement but excluding head movement and remnant movement (i.e. any legal movement in the noun phrase must move the noun, possibly together with dominating structure).[4]

---

[3] A variety of mechanisms to drive movement have been proposed. One idea is that uninterpretable features are part of the lexical representation of some items in certain languages, and such features must be eliminated for convergence; this elimination is accomplished by Internal Merge. Another proposal is that IM is driven by the need to break symmetric or unlabelable configurations, as in Moro's (2000) Dynamic Antisymmetry and Rizzi's (2010) Criterial Freezing. Symmetry-breaking seems particularly ill-suited to explaining neutral order variation, because all of the relevant movements are optional at the cross-linguistic level. Regardless, both proposals apparently fail to achieve Chomsky's goal of linking the duality of semantics to the distinction between EM and IM.

[4] Cinque further assumes Kayne's (1994) Linear Correspondence Axiom (LCA), which necessitates the postulation of additional layers of agreement phrases, interspersed among the nominal categories, to provide landing sites for movement. On this point, see Abels & Neeleman (2012), who argue that the LCA is an unneeded extra stipulation, and the relevant constraint is simply that movement must be to the left.

Cinque's analysis captures important facts: not just the possible and impossible nominal orders,[5] but their derivation as well, hence their bracketed structure. Any purported improvement on this account must preserve both of these descriptive successes, while either capturing additional empirical facts, or simplifying the theoretical apparatus.

It turns out that the same array of orders, and their associated hierarchical bracketed structure, admit a method of generation that appears simpler than Cinque's account (or that of Abels & Neeleman 2012, Steddy & Samek-Lodovici 2011, or related analyses.[6]) This method involves free generation of *n*-ary branching structure[7] superimposed on an arbitrary string of formatives, closely following Chomsky's assertion that Merge applies freely. Yet the account (i) generates all and only the attested orders and appropriate bracketed structures, and moreover (ii) once the bracketing is fixed in any of the legal ways, the assignment of (relative) base hierarchy to the elements follows uniquely. This is an unexpected result, but notable in its simplicity. Here is the procedure:

(1) *Generative procedure over strings*

a. Start with a string of unidentified formatives.

$\mathbf{x\ x\ x\ x}$

b. Place a left bracket just before each formative.

$\mathbf{[}x\ \mathbf{[}x\ \mathbf{[}x\ \mathbf{[}x$

c. Place a matching number of right brackets to form a *legal bracketing*.[8]

$[x\mathbf{]}\ [x\ [x\mathbf{]}\ [x\mathbf{]}\ \mathbf{]}$

d. Scan the bracketed string *right-to-left*, indexing the right brackets in increasing order.

$[x]_{\mathbf{4}}\ [x\ [x]_{\mathbf{3}}\ [x]_{\mathbf{2}}\ ]_{\mathbf{1}}$

e. Copy the index of each right bracket onto the formative immediately following its corresponding left bracket.

$[x_{\mathbf{4}}\ ]_4\ [x_{\mathbf{1}}\ [x_{\mathbf{3}}\ ]_3\ [x_{\mathbf{2}}\ ]_2\ ]_1$

The numbering matches the relative hierarchy of the formatives (see below for nuances), and the bracketed structure is their correct surface structure bracketing. In this case, we derive (2):

---

[5] See, however, Dryer (2018), who offers a different assessment of the typological facts, allowing some orders Cinque (2005) excludes, and explaining the pattern in quite a different way. The present account assumes Cinque's typology is accurate.

[6] It is, however, closely related to the stack-sorting account described in Medeiros (2018). However, while that work commits to a performance-level implementation of the phenomenon, the present formulation keeps to a strictly competence-level generative account.

[7] The *n*-ary branching structure in question, represented as brackets, in effect gives a tree, differing somewhat from the mathematically more sparse set-theoretic conception of Merge. Specifically, a tree has fixed linear order of daughter branches; put another way, this version of External Merge produces an ordered tuple of its operands. This would seem to lose the competition for mathematical simplicity with set-based Merge. But from the point of view of the evolution of the language capacity, simplicity is only a relevant metric if the aspect of language in question is novel. If instead that aspect of language relies on some capacity that is already available in our genetic inheritance, its elegance is not at issue. Allowing linear order to play a role in internal representations of language seems to me not unreasonable, because other animals have the capacity to attend to and internally represent linear order.

[8] Scanning left-to-right, the running total of right brackets cannot exceed the running total of left brackets, and the totals are equal at the end of the string. Strings made up of parentheses grouped in this way are called *Dyck words*; the number of such words of each length forms the famous Catalan number sequence (1, 2, 5, 14, 42, 120...).

(2) [4] [1 [3] [2]]

The simple procedure[9] in (1) generates all and only the attested noun phrase word orders, and their bracketed structure. Importantly, this does not simply repackage the Cinque-style Merge & Move account. In particular, identifying Merge with brackets (one pair of brackets represents the Merge of what the brackets enclose), there is a fixed number of such operations fior all orders: exactly n for n formatives. In a standard framework employing External Merge and Internal Merge, for the same lexical input there are n-1 External Merges, and variable k Internal Merges. Thus, the present account finds a level of uniformity in the generation of the attested orders. Interestingly, this perspective also dissolves the question of "what drives movement": the various attested orders are simply the base-generable structures. There is no notion of steps of movement, and no need to explain them[10]. Conversely, unattested orders are not ruled out by constraints on movement, but simply correspond to impossible bracketings; see below.

Note that no binarity constraint is in force here: brackets may enclose singletons, triples, and beyond, effectively permitting n-ary branching. The placement of left brackets before each surface element, and nowhere else, also departs from standard practice; linguists would expect to be able to have something like [[*ab*]*c*] as a possible structure, but that is ruled out here. This does not mean that "left-branching" structure is impossible. Rather, structure traditionally analyzed as left-branching maps to a horizontal relation between nodes, while right-branching structure comes out as a vertical relation among nodes.[11] While this formulation departs from the usual way of thinking about brackets and their relation to lexical elements, it yields exactly the right orders and their proper structure at a stroke.

For completeness, the following table shows all possibilities with four string formatives.

brackets	formatives	index right brackets	index formatives	order	nominal order
(((())))	(x(x(x(x))))	$(x(x(x(x)_4)_3)_2)_1$	$(x_1(x_2(x_3(x_4)_4)_3)_2)_1$	1234	Dem-Num-Adj-N
((()()))	(x(x(x)(x)))	$(x(x(x)_4(x)_3)_2)_1$	$(x_1(x_2(x_4)_4(x_3)_3)_2)_1$	1243	Dem-Num-N-Adj
((())())	(x(x(x))(x))	$(x(x(x)_4)_3(x)_2)_1$	$(x_1(x_3(x_4)_4)_3(x_2)_2)_1$	1342	Dem-Adj-N-Num
((()))()	(x(x(x)))(x)	$(x(x(x)_4)_3)_2(x)_1$	$(x_2(x_3(x_4)_4)_3)_2(x_1)_1$	2341	Num-Adj-N-Dem
(()(()))	(x(x)(x(x)))	$(x(x)_4(x(x)_3)_2)_1$	$(x_1(x_4)_4(x_2(x_3)_3)_2)_1$	1423	Dem-N-Num-Adj
(()()())	(x(x)(x)(x))	$(x(x)_4(x)_3(x)_2)_1$	$(x_1(x_4)_4(x_3)_3(x_2)_2)_1$	1432	Dem-N-Adj-Num
(()())()	(x(x)(x))(x)	$(x(x)_4(x)_3)_2(x)_1$	$(x_2(x_4)_4(x_3)_3)_2(x_1)_1$	2431	Num-N-Adj-Dem
(())(())	(x(x))(x(x))	$(x(x)_4)_3(x(x)_2)_1$	$(x_3(x_4)_4)_3(x_1(x_2)_2)_1$	3412	Adj-N-Dem-Num
(())()()	(x(x))(x)(x)	$(x(x)_4)_3(x)_2(x)_1$	$(x_3(x_4)_4)_3(x_2)_2(x_1)_1$	3421	Adj-N-Num-Dem
()((()))	(x)(x(x(x)))	$(x)_4(x(x(x)_3)_2)_1$	$(x_4)_4(x_1(x_2(x_3)_3)_2)_1$	4123	N-Dem-Num-Adj
()(()())	(x)(x(x)(x))	$(x)_4(x(x)_3(x)_2)_1$	$(x_4)_4(x_1(x_3)_3(x_2)_2)_1$	4132	N-Dem-Adj-Num
()(())()	(x)(x(x))(x)	$(x)_4(x(x)_3)_2(x)_1$	$(x_4)_4(x_2(x_3)_3)_2(x_1)_1$	4231	N-Num-Adj-Dem
()()(())	(x)(x)(x(x))	$(x)_4(x)_3(x(x)_2)_1$	$(x_4)_4(x_3)_3(x_1(x_2)_2)_1$	4312	N-Adj-Dem-Num
()()()()	(x)(x)(x)(x)	$(x)_4(x)_3(x)_2(x)_1$	$(x_4)_4(x_3)_3(x_2)_2(x_1)_1$	4321	N-Adj-Num-Dem

Table 1: From free bracketing to possible word orders. Left to right, columns show: legal bracketings; with formatives included; with right brackets indexed; with formatives labeled to match the index of their

[9] In fact, we could simplify it further, identifying left brackets themselves with surface positions. Moreover, the procedure described here can equivalently be formulated as follows: (i) generate the full set of Dyck *trees*; (ii) assign hierarchy in bottom-up order (e.g., in the Universal 20 case, <N, Adj, Num, Dem>) to nodes visited in post-order traversal; and (iii) read off linear order by visiting the nodes by pre-order traversal.

[10] This also means that we lose any obvious syntax-internal explanation for the relative typological frequency of different orders (for example, the harmonic orders N-Adj-Num-Dem and Dem-Num-Adj-N are the most common), a matter which Cinque (2005) discusses in detail.

[11] An important question for future research is whether the fundamental branching asymmetry between X-Y and Y-X orders predicted here can be aligned with Wagner's (2005) observations about prosodic asymmetries correlating with relative linear order of predicates and arguments, and of modifiers and heads.

immediate left bracket; as a hierarchical numbered surface form; and exemplified as a noun phrase order. These are exactly the attested orders, according to Cinque (2005).

4 A closer look at the details

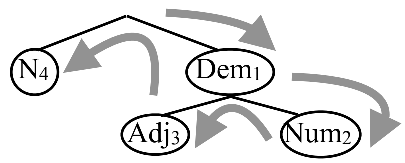
In this section, I explore selected aspects of the account in greater depth. This includes a closer look at how the procedure described above corresponds to standard tree traversal algorithms, how the brackets for nominal orders correspond to Cinque's derivations, and how the account excludes unattested orders.

4.1 The procedure in terms of tree traversals

As mentioned already, the simple procedure over bracketed representations equates, in terms of tree representations, to hierarchization (i.e., labeling) by postorder traversal, and linearization by preorder traversal. While this can be operationalized in string terms, it is helpful to think about the tree format.

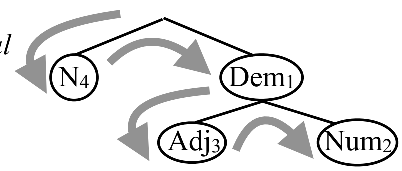
Post-order traversal visits nodes in the tree top-down, and right to left. To illustrate, below is the diagram for 4132 nominal order (N-Dem-Adj-Num) in tree form. The direction of postorder traversal is indicated by the large grey arrows; the subscipt indices reflect the order in which the nodes are visited.

(3) *Post-order traversal*



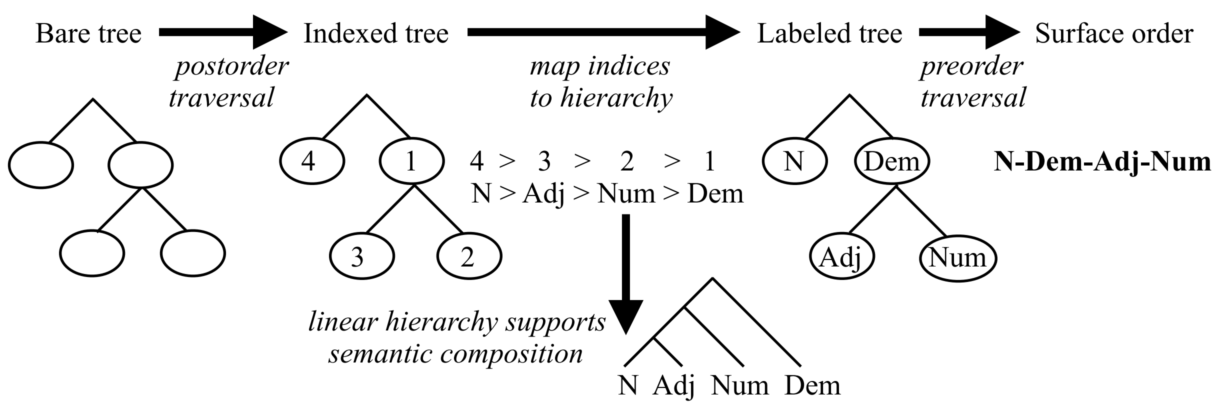
Once the tree has been hierarchized in this fashion, linear order is read off the tree by preorder traversal, which goes top down, left-to-right. The diagram below shows the path of preorder traversal with grey arrows. This path visits the nodes in surface order: N-Dem-Adj-Num.

(4) *Pre-order traversal*



Note that the notion of 'tree' here is much closer to the computer science data structure, and unlike traditional syntactic trees. Notably, words are associated with all nodes; there is no analogue of non-terminal nodes in these trees. I return to these tree representations in section 6. The figure below summarizes the action of this generative architecture in terms of trees.

(5) *Generating N-Dem-Adj-Num (4132) order*



In (5), we illustrate how N-Dem-Adj-Num order is generated. First, a bare n-ary branching tree is built, by Merge applying freely. Postorder traversal assigns indices to each node in the tree. These indices are mapped to a universal sequence encoding the compositional structure; in this case, the hierarchy relevant to Universal 20. This yields a tree with lexical labels on nodes. Preorder traversal of the labeled tree gives surface word order; here, N-Dem-Adj-Num. Separately, the linear representation of the syntactic hierarchy supports semantic composition in the familiar, bottom-up order.

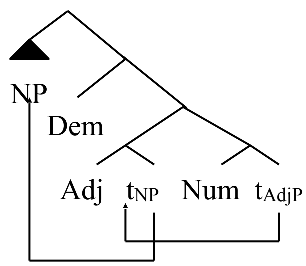
4.2 Correspondence with traditional bracketed representations

Returning to the bracketed string representations, the bracketing generated in this account closely matches that in Cinque's derivations. To illustrate the flavor of the correspondence, we continue with the example illustrated above, for *4132* order. Translating this example to the noun phrase hierarchy relevant to Universal 20, 1 = Dem, 2 = Num, 3 = Adj, and 4 = N. Thus, the structure is:

(6) [N] [Dem [Adj] [Num]]

Illustrated below is a (simplified) Cinque-style derivation of this order.

(7)



In this derivation, two movements occur: first of the Adj-NP complex to precede the Num element, followed by subextraction of the NP and movement to a specifier position before Dem.[12] The following bracketed expression represents the resulting structure:

(8) [[NP] [Dem [[Adj t] [Num t]]]]

If we highlight only the bracketed pairs where the left bracket immediately precedes a lexical element (supposing this is true for the NP as well, i.e. NP ~ [N]):

(9) [**[**N**]** **[**Dem [**[**Adj t**]** **[**Num t**]**] **]**]

Keeping only such "lexically headed" bracket pairs, we find exactly the bracketing derived under the direct generative procedure described in this paper:

(10) [N] [Dem [Adj] [Num]]

4.3 Unattested orders require impossible bracketing

Consider in more detail how the unattested orders are ruled out. With a hierarchy of just three elements (say, Dem=1, Adj=2, N=3), there are six logically possible word orders, and of these five are attested as neutral noun phrase orders. One permutation, *213 (*Adj-Dem-N), does not occur as a basic noun phrase order. The system described here explains this systematic gap.

---

[12] This example is chosen to illustrate that the present account can provide an analysis even of somewhat exotic attested orders. A notable property of this configuration is that the surface order contains cross-serial selectional dependencies among elements of the hierarchy; see section 6 below.

Given that left brackets occur immediately before each surface element, and nowhere else, we can begin to fill in what a 213 order would have to look like as a bracketed string.

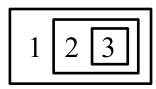
(11) $[_2\ 2 \ldots [_1\ 1 \ldots [_3\ 3 \ldots$

Because right brackets are numbered right-to-left, they must occur in the sequence $]_3 \ldots]_2 \ldots]_1$. Furthermore, right brackets must follow the left bracket and element they match. Therefore, the entire sequence of right brackets must follow the element 3. This gives us:

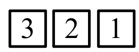
(12) $[_2\ 2\ [_1\ 1\ [_3\ 3\ ]_3\ ]_2\ ]_1$

This is not a legal bracketing; the boundaries of bracketing 1 and 2 cross. To clarify this point, we can think of brackets as denoting the edges of “boxes”. In generated orders, any pair of boxes may be in a containment relation, or be disjoint; they cannot overlap partially. Illustrating with 123 and 321 order and appropriate bracketing:

(13) $[_1\ 1\ [_2\ 2\ [_3\ 3\ ]_3\ ]_2\ ]_1$

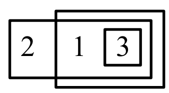


(14) $[_3\ 3\ ]_3\ [_2\ 2\ ]_2\ [_1\ 1\ ]_1$



But the unattested 213 order entails overlapping boxes:

(15) $[_2\ 2\ [_1\ 1\ [_3\ 3\ ]_3\ ]_2\ ]_1$.



Given the logical flow implicit in this procedure - first bracket an undifferentiated string, and the relative hierarchy of word order positions follows - the unattested 213 order cannot be generated. Instead, the relevant bracketing must correspond to a 123 order; bracketing determines hierarchy.

5.0 Generating the Final-Over-Final Condition

A crucial aspect of the explanation of Universal 20 here is the particular way in which the nominal hierarchy is mapped to freely-generated branching structure. This includes not just the choice of post-order traversal, one of several standard tree traversal algorithms,[13] but furthermore exactly how to compress linguistic hierarchical relations into a consistent linear order that can be mapped to the sequence of nodes visited. In this regard, it is notable that fixed relations among syntactic elements seem to come in (at least[14]) two flavors: roughly, selection and adjunction, or head-complement and head-adjunct relations.

Recall that post-order traversal visits nodes/right brackets "outside-in", and from right to left. It is natural to assign numerical indices in the same order: the outermost right bracket/node is 1, the next node visited by postorder traversal is 2, etc. We define the hierarchical ordering relation '>' in the usual way with respect to this indexing of the traversal sequence; for example, $2 > 1$.

In these terms, suppose a head H and adjunct A are mapped to this sequence such that $H > A$. That means the head is effectively "deeper" in the hierarchy (here, reached later by post-order traversal of the tree, and thus having a greater numerical index) than its adjunct, a familiar analysis.

Moreover, if H has several adjuncts A_2, A_1, with A_2 the "closest" in traditional representations, then we will have $H > A_2 > A_1$. Restricting attention to a hierarchy comprised of a head and a series of adjuncts to that head, we will find *213-avoidance: *A_2-A_1-H . This is the pattern seen in Cinque's

[13] In fact, the bracket-numbering scheme here, together with the standard method of reading linear order, generates what Feil *et al* (2005) call the *PostPre permutations*: labeling trees by post-order traversal, and reading out the labeled nodes by pre-order traversal.

[14] Whether these two modes suffice to model conjunction, or if further complications are necessary, is another important question set aside here.

version of Universal 20, and arguably in verb clusters (though see Salzmann 2019 on attested 213 verb clusters, and Abels 2016 on the appearance of Universal 20 effects in other domains, including among verbs and their arguments).

What about the other kind of relation, between heads and their complements? In the traditional analysis, heads and complements are in a symmetric hierarchical relationship. The present context provides no basis for such a symmetry, and we must make a choice: heads must be hierarchically above, or below, their complements. That is, we are mapping the syntactic hierarchy onto the tree traversal sequence, which is necessarily linear.

Suppose that head-complement relations obey the same *H > other* convention: head H and complement C are mapped to the post-order traversal index sequence such that H > C[15]. This will produce the basic phenomenology of the Final-Over-Final Condition (FOFC; Sheehan *et al* 2017) in structures characterized by head-complement relations, a happy result.

To see this, consider a configuration with nested complementation: head H_a takes a complement headed by H_b, which in turn has complement C. The hierarchical order relation is then $H_a > H_b > C$. The forbidden permutation is *H_b-C-H_a. That banned order is traditionally described as a head-final phrase (H_aP) dominating a head-initial phrase (H_bP); this is exactly the configuration ruled out by FOFC. For example, if we have head Aux taking a complement headed by V, with its own complement Obj, the hierarchy is Aux > V > Obj. We correctly exclude the unattested *213 order *V-Obj-Aux. Since the reasoning is about heads and complements (rather than just verbs and auxiliaries), we expect this ordering constraint to generalize to any partial hierarchy characterized by head-complement relations, reconstructing the core predictions of FOFC.

What about structures with both adjuncts and complements? Sheehan (2017) argues that FOFC extends to certain adjunct relations; for example, adverbials obey the FOFC generalization. Concretely, parallel to the FOFC effect *V-Obj-Aux, *V-Adv-Aux is unattested. A full discussion is put aside, but note that the ordering effect in question is correctly predicted by this system. This follows from the already assumed underlying hierarchical sequence, Aux > V > Adv; unattested *V-Adv-Aux is the forbidden *213 permutation.

In existing models of syntactic combination, complements are the closest element to the head, the first-merged modifier. Adjuncts are further away (within or Chomsky-adjoined to the same phrase, or introduced by further functional superstructure). Essentially the same relation is encoded by the present ordering, H > Comp > Adjunct (*i.e.*, the complement is the unique closest element to the head). However, in the usual understanding, while the H-adjunct relation resolves structurally as one involving asymmetric hierarchy (the adjunct is "above" the head, on either theory of adjuncts), the head-complement relation is famously, and problematically, symmetric. The present approach avoids this unwanted symmetry (admittedly, by stipulating one choice of head-complement hierarchical ordering), with promising consequences for word order constraints.

6.0 Generating some well-known crossing dependencies

Bresnan *et al* (1982) discuss unbounded crossing subject-verb dependencies in Dutch (there is also a long-distance crossing dependency between object and the main verb here, shown in grey).

(16) ...omdat ik Cecilia Henk de nijlpaarden zag helpen voeren (from Steedman 2000: 25)
...because I Cecilia Henk the hippos saw help feed
'...because I saw Cecilia help Henk feed the hippos'

[15] While the H > Adjunct hierarchical relation can be rationalized in terms of traditional analyses, breaking the head-complement symmetry in this particular way is, for now, simply a stipulation. However, this stipulation has appealing consequences.

Shieber (1985) discusses similar facts in Swiss German, which furthermore exhibits long-distance cross-serial case dependencies. Interestingly, with the system established this far, we can base-generate these orders[16]. I assume the example above contains (at least) the categories shown below. The internally complex object *de nijlpaarden* 'the hippos' is treated as a single unit at this level. I segment a Tense suffix from inflected and non-finite verbs, even if realized as zero.

(17) ...omdat ik Cecilia Henk de nijlpaarden zag-0 help-en voer-en
C S_1 S_2 S_3 O_3 V_1 T_1 V_2 T_2 V_3 T_3

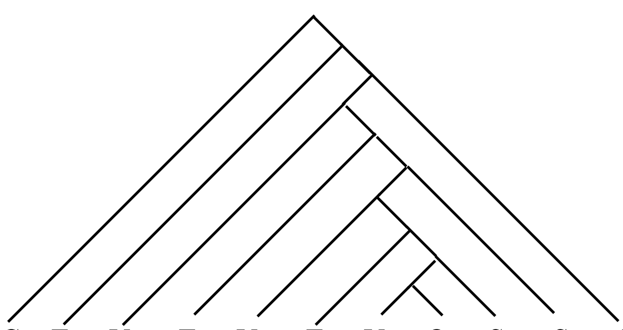
The categories in this complex sentence will be rendered as a single linear hierarchy, which we assemble incrementally. Recall that heads get a higher index than all their arguments and adjuncts: H > Arg; H > Adj. Where a head takes both arguments and adjuncts, I assume the relative hierarchy is H > Arg > Adv. If there are multiple arguments of a head, the complement is closest to the head: H > Comp > Arg'.

Concretely, for verb head V and complement object O, V > O. The same hierarchy holds for a verb and complement clause: V > CP. A ditransitive verb would have V > DO > IO. If there is an adverbial and an object, the hierarchy is V > O > Adv.

Adding the layer for Tense and subject, the order is T > V > O > S. To align with modern proposals, we might include 'little *v*' between T and V, and group the subject with it: T > (*v* > V > O > S). Note, though, that the parentheses above are for visual emphasis only. Since no overt morpheme obviously realizes little *v*, I omit it from the representations here. If a complementizer is present, I assume it takes TP as complement: C > T > V > O > S.

We are now in a position to integrate the relative order of verb and complement clause with the internal clause order just elaborated[17]. For a single layer of clausal embedding, $[_{CP} \ldots [_{CP}\underline{\quad}]\,]$, we have the following: $C > T_1 > V_1 > \underline{T_2 > V_2 > O_2 > S_2} > S_1$. Now replacing the O_2 position with one more embedded clause, we arrive at our unified order for three clauses, which matches the Dutch example sentence above. For clarity, I also write the postorder indices aligned to this linear hierarchy, and include a superposed tree to show how the structure composes semantically. The reader can verify that composition proceeds in the usual bottom-up fashion, starting with the most deeply embedded verb-object pair.

(18) *Integrated hierarchical order for sentence (16) with postorder index and composition tree*



$C > T_1 > V_1 > T_2 > V_2 > T_3 > V_3 > O_3 > S_3 > S_2 > S_1$ Hierarchical order
11 10 9 8 7 6 5 4 3 2 1 Postorder index

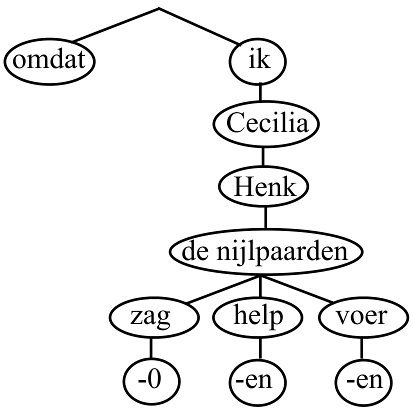
[16] However, see Stabler (2004) for the conclusion that there are four different varieties of cross-serial dependency constructions, with differing properties. I restrict attention to the selection of examples in this section.

[17] At least for these structures, note that we are implicitly developing a straightforward account of recursion (by substitution). I leave a more complete consideration of recursion in other domains to future research.

Given this mapping from the syntactic hierarchy to the post-order index sequence, we can easily recover the tree structure corresponding to the Dutch surface order[18], shown in (19).

(19) ...omdat ik Cecilia Henk de nijlpaarden zag-0 help-en voer-en

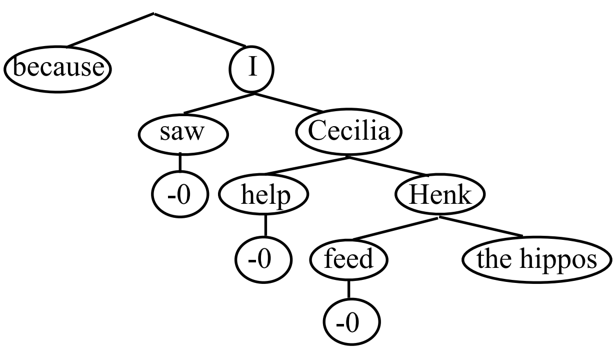
Category	C	S_1	S_2	S_3	O_3	V_1	T_1	V_2	T_2	V_3	T_3
Index	11	1	2	3	4	9	10	7	8	5	6
Brackets	( )	(	(	(	(	(	( ))	(	( ))	(	( ))))))



Having resolved the relevant syntactic hierarchy as a universal linear sequence, we can readily construct the same kind of bracketed derivation for other orders of the same elements. The tree for the English version of this sentence is in (20).

(20) ...because I saw -0 Cecilia help -0 Henk feed -0 the hippos

Category	C	S_1	V_1	T_1	S_2	V_2	T_2	S_3	V_3	T_3	O_3
Index	11	1	9	10	2	7	8	3	5	6	4
Brackets	( )	(	(	( ))	(	(	( ))	(	(	( ))	( ))))



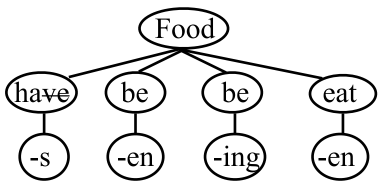
Finally, note that the present system can generate the more limited pattern of crossing dependencies that arises in English Affix-Hopping (Chomsky 1957), as seen in the following example.

(21) Food ha-s be-en be-ing eat-en

[18] An important question is whether these structures provide a basis for a successful theory of prosody. While there are some promising hints in the close correspondence of trees derived in this theory to, for example, Cinque's derivations of nominal orders, as well as the somewhat "flatter" shape of certain trees here, I leave this question aside for future work.

As Chomsky pointed out, the verbal affixes group with the preceding auxiliary in their distribution and meaning, despite being separated in surface order by an intervening verb. To accommodate this pattern, suppose that an auxiliary Aux and associated displaced affix -Fx, is resolved with hierarchical order Aux > -Fx > VP-complement[19]. This order is indeed generated by the present account, with structure as shown in (21)[20].

(21)	Food	ha~~ve~~	-s	be	-en	be	-ing	eat	-en
Cat.	Obj	Aux$_1$	T	Aux$_2$	-Fx$_1$	Aux$_3$	-Fx$_2$	V	-Fx$_3$
Index	1	8	9	6	7	4	5	2	3
Brackets	(	(	( ))	(	( ))	(	( ))	(	( )))



7.0 Conclusion

Implementing Merge as an operation building bare ordered trees, lexicalized and linearized by traversal algorithms, we predict a striking array of word order constraints. These predictions align well with current empirical descriptions, even offering simple analyses of exotic constructions exhibiting cross-serial dependencies. In this view, there is nothing "extra" or stipulated about the kind of displacement involved in deriving neutral orders; the full set of typologically possible orders is base-generated. This unification of movement with fundamental structure-building goes further than Chomsky's framing of movement as Internal Merge, where Internal Merge involves one or more additional operations over and above the constant number of External Merges required to join the lexical items involved. Not so here, where exactly the same number of External Merge operations (*i.e.*, pairs of matched brackets) is involved for all neutral orders: exactly *n* such for *n* items.

That said, there is clearly still a need for movement in the present framework, over and above the base generation of neutral orders: effects like *wh*-movement and topic and focus displacement produce other orders.[21] In that sense, we have not eliminated the need for Internal Merge or some similar mechanism to effect the needed movements, and so have not simplified this aspect of the account. Notice, though, that the residue of actual movements under this account is precisely the set of non-information-neutral transformations. This result aligns with Chomsky's suggestion that the duality of semantics is tied to the distinction between External Merge and Internal Merge: EM builds the base thematic structure, and IM induces discourse-information effects. This result fails to obtain on the standard view that IM is involved in the derivation of information-neutral orders, as for example in Cinque's account of Universal 20. But in the system presented in this paper, External Merge alone (the generative procedure) produces the full array of typologically available neutral orders, and Internal Merge (real movement, however implemented) applies only semantically contentful transformations. That seems the proper cut.

Of course, the theory developed here is only a fragment. I have not demonstrated how this system generalizes to word order in all domains, nor attempted to spell out how "real" movement works, nor accounted for any number of important grammatical phenomena such as coordination, ellipsis, binding,

[19] One way of understanding this proposed hierarchy is to say that the affix (say, *-ing*) is a head sandwiched between its selecting auxiliary (*be*) and the host verb. Or we might prefer an analysis where the auxiliary and associated affix are treated as a unit for interpretation, mirroring Chomsky's (1957) analysis, where the two were introduced as a single unit (*e.g.*, *be+ing*) from the lexicon.

[20] It's not clear if the passive movement of the object should be base generated, or if it is obligatorily an instance of "real" movement. It is at least possible to generate it with just this mechanism.

[21] Thanks to David Adger for discussion on this point.

agreement, and so on. These are all important topics, and much more work will be required to determine if they might find satisfying accounts within this framework.

References:

Abels, K., & Neeleman, A. (2012). Linear asymmetries and the LCA. *Syntax*, 15(1), 25-74.

Abels, K. (2016). The fundamental left–right asymmetry in the Germanic verb cluster. *The Journal of Comparative Germanic Linguistics* 19(3), 179-220.

Bresnan, J., Kaplan, R.M., Peters, S., & Zaenen, A. (1982). Cross-serial dependencies in Dutch. *Linguistic Inquiry* 13(4), 613-635.

Chomsky, N. (1957) *Syntactic Structures*. The Hague: Mouton.

Chomsky, N. (2007). Approaching UG from Below. In Hans-Martin Gärtner, H.-M., & Sauerland, U. (eds.). *Interfaces + Recursion = Language? Chomsky's Minimalism and the View from Syntax-Semantics.* Studies in Generative Grammar. Berlin: Mouton de Gruyter.

Cinque, G. (1999). *Adverbs and Functional Heads: A Cross-linguistic Perspective*. New York: Oxford University Press.

Cinque, G. (2005). Deriving Greenberg's universal 20 and its exceptions. *Linguistic Inquiry* 36(3), 315-332.

Dryer, M. (2018). On the order of demonstrative, numeral, adjective, and noun. *Language* 94(4), 798-833.

Feil, T., Hutson, K. and Kretchmar, R.M. (2005). Tree traversals and permutations. *Congressus Numerantium* 172, 201-221.

Greenberg, J. (1963). Some universals of grammar with particular reference to the order of meaningful elements. In Greenberg, J., (ed.), *Universals of language*, 73-113. Cambridge, MA: MIT Press.

Holmberg, A. (2000). Deriving OV order in Finnish. in Svenonius, P., (ed.), *The Derivation of VO and OV*. Philadelphia: John Benjamins.

Kayne, R. (1994). *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.

Medeiros, D. (2018). ULTRA: Universal Grammar as a Universal Parser. *Frontiers In Psychology*. Ed. Gallego, A.

Moro, A. (2000). *Dynamic Antisymmetry*. Cambridge, MA: MIT Press.

Rizzi, L. (1997). The fine structure of the left periphery. In Haegeman, L. (ed.), *Elements of Grammar: Handbook of Generative Syntax*. Dordrecht: Kluwer.

Rizzi, L. (2007). On Some Properties of Criterial Freezing. *CISCL Working Papers on Language and Cognition* 1: 145-158.

Salzmann, M. (2019). On the limits of variation in Continental West-Germanic verb clusters. Evidence from VP-stranding, extraposition and displaced morphology for the existence of clusters with 213 order. *Journal of Comparative Germanic Linguistics* 22, 55-108.

Sheehan, M., Biberauer, T., Roberts, I., & Holmberg, A. (eds.). (2017). *The Final-over-Final Condition: A syntactic universal* (Vol. 76). Cambridge, MA: MIT Press.

Sheehan, M. (2017). The final-over-final condition and adverbs. in Sheehan, M., Biberauer, T., Roberts, I., & Holmberg, A. (eds.) *The Final-over-Final Condition: A syntactic universal* (Vol 76), 97-120. Cambridge, MA: MIT Press.

Shieber, S.M. (1985). Evidence agaisnt the context-freeness of natural language. *Linguistics & Philosophy* 8(3), 333-344.

Stabler, E. (2004). Varieties of crossing depemdencies: structure dependence and mild context sensitivity. *Cognitive Science* 28: 699-720

Steddy, S. & Samek-Lodovici, V. (2011). On the ungrammaticality of remnant movement in the derivation Greenberg's universal 20. *Linguistic Inquiry* 42(3), 445-469.

Steedman, M. (2000). *The syntactic process*. (Vol. 24) Cambridge, MA: MIT press.

Wagner, M. (2005). Prosody and recursion. Ph.D. dissertation, MIT.